



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

RELIABILITY OF GRADING WORK IN HISTORY

DANIEL STARCH AND EDWARD C. ELLIOTT
University of Wisconsin

This is the third of a series of studies on the variability of values placed by teachers upon examination papers. The first two dealt with English and mathematics respectively.¹ The present investigation, dealing with work in history, was conducted in exactly the same manner as the two preceding ones. An answer paper written as a final examination in United States history in a large high school in Wisconsin was manifolded by plates so as to reproduce the original precisely as presented by the pupil.

QUESTIONS

Write on any five of the seven questions.

1. Contrast the motives and methods of settlement of the French and English colonists in America.
2. (a) Point out all the possible points of difference between what the English and American idea of representative government was about 1775. (b) Give two concrete illustrations of how the above difference caused open friction.
3. (a) Explain clearly what the British plan of attack for 1777 was. (May be outlined.) (b) Point out why the date, October 17, 1777, was such an important date in American history.
4. Describe the "Period of Confederation." (a) Name of instrument of government used. (b) Defects in plan as proved by experience.
5. (a) Trace the steps leading up to the Federal Convention. (b) What objections were given to the ratification of the constitution? (c) Why is the constitution considered such a "wonderful instrument of government"?
6. (a) Contrast the personal characteristics and political policies of Hamilton and Jefferson. (b) What was Washington's policy relative to foreign alliances and how has it been observed since that time?
7. Describe industrial conditions in the United States in 1800 as to: (a) various industries engaged in; (b) their relative importance and why. (c) Define the various kinds of tariff. (d) Why did the South object to the protective tariff?

¹D. Starch and E. C. Elliott, *School Review*, XX, 442-57; XXI, 254-59.

History

I The English settlers that came to America, came generally for wealth, or on account of religious persecutions. The first English settlers were in search of wealth which they failed to find along the American coasts. The French came for adventure and with some attempt of settlement. They gain the friendliness of the Indians from the start but some of the English tried to drive the native inhabitants before them as they entered the new land and traveled westward. The French engaged almost from the start in fur trading and in exploring expeditions thru Canada. The English were the downfallen nobles who came to this country in hopes of regaining their wealth. They knew nothing about agriculture and therefore starved to death or returned to England. The French landed in the warmer months prepared for exploration and adventure which result in better settlements at first. The English later on adopted conditions favorable to American settlement.

[Reduced to about one-half]

ANSWERS¹

1. The English settlers that came to America, came generally for wealth, or on account of religious persecutions. The first English settlers were in search of wealth which they failed to find along the American coasts. The French came for adventure and with some attempt of settlement. They gain the friendliness of the Indians from the start but some of the English tried to drive the native inhabitants before them as they entered the new land and traveled westward. The French engaged almost from the start in fur trading and in exploring expeditions thru Canada. The English were the downfallen nobles who came to this country in hopes of regaining their wealth. They knew nothing about agriculture and therefore starved to death or returned to England. The French landed in warmer months prepared for explorations and adventure which resulted in better settlements at first. The English later on adopted conditions favorable to American settlement.

2. The English thought the king was the representative of the people and who was to have all power. The American idea was of a representative government was a government with someone at the head with the power in the hands of the people or some representatives representing the people. That is, the passage of the stamp act by England, which was a tax upon the American people without representation. The English government simply passed the act and put it into effect without asking the American people whether or not they thought it just. Another instance was the tax levied upon tea imported into American colonies.

3. The British plan of attack for 1777 was to separate the New England states from the other states by gaining control of the Hudson River and the Lake Champlain region. This was to be done by sending Howe to Philadelphia who was to receive reinforcement from Clinton left at New York. Burgoyne was to come down thru the Lake Champlain region and receive help from Carlton of Canada and from St. Leger by way of the Mohawk River. The expedition failed because of the few men left with Clinton, the jealousy of Carlton and the defeat of St. Leger. (b) The date Oct. 17, 1777 is important in Am. history because it is the date beginning the independence of the United States and also setting forth capability of the United States in defending herself against other nations.

6. (a) Jefferson was a man who did not look at the showy side of life the way Hamilton did. His dress compared with that of Hamilton was poor, taking under consideration the offices held by Jefferson. Jefferson rode horseback to congress and tied his own horse while Hamilton thought he should of put on more style. Jefferson's idea was peace with all nations. He also sympathized with the French which were revolting at this time. Hamilton although a good man in politics seemed to adopt unjust methods in bringing forth his ways. Hamilton was a Federalist and Jefferson a republican. During one election Hamilton was among the men of the Federalist party who

¹ The errors are reproduced as in the original.

planned that in throwing away their votes would bring their party representatives into offices. Their plan were defeated by themselves throwing too many votes away.

b) Washingtons policy was peace with all nations, and an independent nation on equal terms with all nations. This policy has be followed out by presidents following him.

7. The industries engaged in, in the U. S. in 1800 are agriculture, commerce, and fishing. Commerce was about this time beginning to take an important stand in industrial conditions in U. S. Agriculture was important because it had been the only occupation of the American people and cotton was their chief export. Fishing was of somewhat importance but not so much as agriculture and commerce.

(c) The various tariffs are tariffs for protection and revenue. Tariff for protection is a duty levied upon goods imported into the country to bring the outside manufactured articles up to a higher price than the articles manufacture in the U. S. Tariff for revenue is a duty on imports or exports as a tax to help pay government expenses.

(d) The south objected to protective tariff because it meant exportation of their cotton to Am. ports which paid less, the journey was more dangerous, and the north would be reaping all the benefits. The overlan route to the northern states was in such a condition it could not be traveled upon. The water routs contain dangerous points and the south would have to sell to the northern ports cheaper than to England and then pay a higher price for the manufactured article.

A set of questions and a copy of the answer paper were sent to some two hundred high schools in the Middle West with the request that the principal teacher of history grade this paper according to the practices and standards of the school.

One hundred and twenty-two papers were returned. Eight could not be used because the data were incomplete. Seventy were returned from schools whose passing grade was 75, twenty from schools whose passing grade was 70, twenty from schools whose passing grade was 80, and four from schools whose passing grade was 65. The comments and criticisms on the returned papers show that they were evaluated with much care and discrimination.

The values assigned by the seventy schools whose passing grade was 75 are shown in the distribution chart of Fig. 1. The range of the grades is indicated along the base line and the number of schools assigning a g'iven grade is indicated by the number of dots above that grade. Thus the grade 70 was given to the paper by six

different schools and the grade 71 by three different schools, etc. The distribution of these marks is very similar to that found for the English and mathematics papers. The extreme range extends from 43 to 92. The median value is 70.8 and the probable error is 7.7.

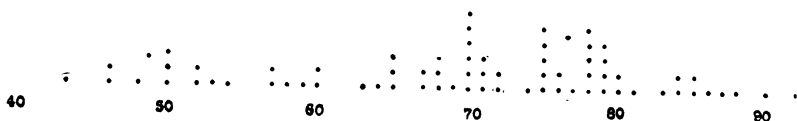


FIG. 1

The values assigned by the twenty schools whose passing grade was 80 are 70, 70, 83, 84, 80, 75, 71, 85, 62, 50, 53, 65, 80, 76, 75, 72, 55, 75, 78, 75. The median grade is 74.8. The values assigned by the twenty schools whose passing grade was 75 are 45, 60, 51, 65, 72, 75, 65, 63, 61, 18, 35, 88, 77, 77, 48, 66, 70, 67, 67, 55. The median grade is 65. The four schools whose passing grade was 65 returned marks of 66, 40, 76, and 52.

The chief results of this series of investigations may be summarized as follows:

1. The marks assigned to the same paper by different teachers vary enormously, in fact much more widely than the average teacher would anticipate. The findings with the history paper fully corroborate the findings with the English and mathematics papers. The range and distribution of the marks of papers in these three subjects are almost identical. The extremes in each case extend nearly over the entire marking scale.

2. The variability or unreliability of marks is as great in one subject as in another. Contrary to current belief, grades in mathematics are as unreliable as grades in language or in history. The probable error is very nearly the same in all subjects, being 5.4 for the English grades, 7.5 for the mathematics grades, and 7.7 for the history grades. Hence the variability of marks is not a function of the subject but a function of the examiner and of the method of examination.

3. The immense variability of marks tends obviously to cast considerable discredit upon the fairness and accuracy of our present methods of evaluating the quality of work in school. No matter how much anyone may wish to minimize the utility of marks,

they have, nevertheless, an indispensable administrative value from the standpoint of the school, and a real personal value from the standpoint of the pupil.

The chaotic status revealed by our present inquiry raises two rather important questions: First, what are the factors that produce such wide divergences in the evaluation of school work? And second, what may be done to secure greater uniformity and more objective reliability?

An answer to the first question has been worked out by means of tests conducted by one of the writers¹ and published elsewhere. These results may be briefly quoted here. There are four major factors that produce the variability of marks: "(1) Differences among the standards of different schools, (2) differences among the standards of different teachers, (3) differences in the relative values placed by different teachers upon various elements in a paper, including content and form, and (4) differences due to the pure inability to distinguish between closely allied degrees of merit." Taking the probable error of 5.4 found for the English grades, the special tests showed that the "fourth factor contributes 2.2 points, the third 2.1 points, the second 1.0 point, and the first practically nothing toward the total" probable error. Hence the largest factors are the fourth, third, and second.

The second problem is more difficult to solve. One suggestion would be the adoption by all schools of some uniform marking system such as is outlined in the article just quoted;² that is, in brief, the adoption of a scale with a definite number of steps and the preparation of a standard curve or table showing the number of times each particular step should in the long run be assigned. Courses in education should give instruction regarding the technique and methods of marking and evaluating school work. Teachers could thus be led to appreciate the problems involved and to make efforts toward greater uniformity.

Another possible suggestion would be the development and general use of standard tests and scales for measuring efficiency in all subjects similar to the ones already devised for arithmetic, composition, and handwriting.

¹D. Starch, "The Reliability and Distribution of Grades," *Science*, XXXVIII, 630-36.